DOCUMENT RESUME

ED 028 799 LI 001 456

Development of a Computer Processing Center for the New England State University Libraries. Final Report.
Inforonics, Inc., Cambridge, Mass.
Spons Agency-Council on Library Resources, Inc., Washington, D.C.; New England Board of Higher Education, Wellesley, Mass.
Report No-CLR-354
Pub Date 13 Jul 67
Note-61p.
EDRS Price MF-$0.50 HC-$3.15
Descriptors-*Automation, *Cataloging, *Centralization, Computer Programs, Library Acquisition, *Library Networks, *Library Technical Processes, Systems Development, *University Libraries
Identifiers-*Machine-Readable Cataloging, MARC

This report describes the system design of a regional computer center for the libraries of New England State Universities. The function of this center is to provide library technical processing service to the participating libraries. These services will include: (1) catalog data file creation and maintenance, (2) catalog data file search and retrieval, (3) production of catalog card sets, (4) production of book labels, (5) production of book pockets, and (6) acquisitions control. The computer will be used a s a tool to provide processing services and will be a conduit for current cataloging information in machine form produced by the Library of Congress Machine-Readable Cataloging (MARC) project. Three tasks were defined which comprise the work required to implement the system: Task 1, Catalog data file creation, Task 2, Catalog data file searching, and Task 3, Acquisitions processing. The programs for these tasks and the machine configurations to run them, both in demonstration and in regional center operation, are discussed in this report. (Author/CC)

ED028799

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

LI 001456

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

DEVELOPMENT OF A COMPUTER PROCESSING CENTER FOR THE NEW ENGLAND STATE UNIVERSITY LIBRARIES

submitted to

The New England Board of Higher Education

and

The Council on Library Resources

ERIC
MAY 29 '69
UNIVERSITY OF MINNESOTA

Final Report
Contract No. CLR-354

LI 001456

July 13, 1967

*new concepts in information organization, processing, and presentation

INFORONICS, INC.
806 MASSACHUSETTS AVENUE, CAMBRIDGE, MASSACHUSETTS 02139, TEL. (617) 547-1750
146 MAIN STREET, MAYNARD, MASSACHUSETTS 01754, TEL. (617) 897-8815
927 15TH STREET, N. W., WASHINGTON, D. C. 20005, TEL. (202) 638-6862

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	1
1.1	STUDY PROGRESS AND OBJECTIVES	2
2.	OVERALL PLAN AND DEVELOPMENT TIMETABLE	4
2.1	TASK 1 SYSTEM: CATALOG DATA FILE CREATION	4
2.1.1	SERVICES PROVIDED BY TASK 1 PROGRAMS	4
2.1.2	THE USE OF TASK 1 SERVICES IN CATALOGING NEW TITLES	6
2.1.3	USE OF TASK 1 SERVICES FOR RECLASSIFICATION	6
2.2	TASK 2 SYSTEM: CATALOG DATA FILE SEARCHING	7
2.2.1	PROPOSED OPERATION	8
2.2.2	REGIONAL CENTER SEARCH OPERATION	8
2.2.2.1	BATCH PROCESSING	8
2.2.3	SEARCH STRATEGY	9
2.2.3.1	SEARCH BY L.C. CARD NUMBER	9
2.2.3.2	TITLE SEARCH	9
2.2.3.3	SEARCH BY AUTHOR	10
2.2.3.4	CONFERENCE PROCEEDINGS	10
2.2.4	USE OF MARC DATA	10
2.2.5	COMPUTER PROGRAMMING REQUIREMENTS	11
2.2.5.1	DIRECTORY GENERATOR	11
2.2.5.2	FILING KEY GENERATOR	11
2.2.5.3	DIRECTORY SEARCH AND PRINTOUT PROGRAM	11
2.2.5.4	DIRECTORY MERGE	11
2.2.6	SYSTEMS CONFIGURATION FOR TASK 2	12

2.3 TASK 3 SYSTEM: ACQUISITIONS, SEARCH, AND CONTROL 13

2.3.1 SERVICES TO BE PROVIDED BY TASK 3 14

2.3.1.1 HOLDINGS SEARCH 14

2.3.1.2 PRODUCTION OF PURCHASE ORDER RECORD 15

2.3.1.3 PRODUCTION OF CONTROL FILE INFORMATION 15

2.3.1.4 BOOK PROCESSING 16

2.3.1.5 MAINTENANCE AND UPDATING OF WANT LISTS TO BE DISTRIBUTED TO ANTIQUARIAN DEALERS 16

2.3.2 DETAILED DESIGN OF AN ACQUISITION SYSTEM 16

2.3.3 PREPARATION OF A REQUEST FOR A BOOK 17

2.3.4 SEARCH 18

2.3.4.1 ON-ORDER FILES 18

2.3.4.2 SEARCH HOLDINGS FILE 19

2.3.5 PURCHASE ORDER CONTROL 19

2.3.6 RECEIPT OF ITEM AT ORIGINATING LIBRARY 19

2.3.7 USE OF COMMON ORDER-PROCESSING PROGRAMS 19

2.3.8 TRANSMITTAL OF BIBLIOGRAPHIC DATA TO CATALOGING 20

2.3.9 COMPUTER PROGRAMMING REQUIREMENTS (TASK 3) 20

2.3.9.1 FILE CREATION PROGRAM 20

2.3.9.2 FILE UPDATE PROGRAM 20

2.3.9.3 FILE SEARCH 20

2.3.9.4 FILE PRINTOUT 20

2.3.9.5 FILE TRANSFER 21

2.3.10 SYSTEM CONFIGURATION AND COST ESTIMATES FOR TASK 3 23

2.4 TASK 4: INITIAL USE OF TELETYPE NETWORK AND DEMONSTRATION OF PROGRAMS AND SERVICES 24

2.5 THE FUTURE NEW ENGLAND LIBRARY NETWORK 24

APPENDIX 1 - THE OVERLAP EXPERIMENT

APPENDIX 2 - NEBHE COMPATIBILITY WITH MARC

APPENDIX 3 - LIBRARY OF CONGRESS CLASSIFICATION SCHEDULES AND THE NEBHE PROJECT

1. INTRODUCTION

This report describes the system design of a regional computer center for the libraries of the New England state universities. The function of this center is to provide library technical processing services to the participating libraries. These services will include:

1. Catalog data file creation and maintenance.
2. Catalog data file search and retrieval.
3. Production of catalog card sets.
4. Production of book labels.
5. Production of book pockets.
6. Acquisitions control.

Although the services listed above are the only ones that have been planned thus far, the services that the regional center might eventually provide need not be restricted to these but might extend to any of the following:

1. Production of accessions lists.
2. Production of book catalogs.
3. Production of serial holdings lists.
4. Circulation control.
5. Serial records control.
6. Maintenance of subject and other authority files.

The computer will be used as a tool to provide processing services, but equally important, it will be a conduit for current cataloging information in machine form produced by the Library of Congress MARC project.

The mechanized regional center will provide service and speed that would not be economically attainable in separated or manual systems. The planned growth rates of the participating libraries (between 10% and 20% per year) are now matched by increasing problems of backlog and a short supply of skilled library personnel. The regional computer center will perform much of the clerical operations that now burden the libraries' professional staff, and will in addition provide better service to the libraries' users by closing the time gap between acquisition and availability.

In this study we have developed a system design of scheduled growth, under which a succession of service functions are to be programmed, tested, and placed in operation. Three tasks have been defined which comprise the systems and programming work required to implement the system.

Task 1: Catalog data file creation.

Task 2: Catalog data file searching.

Task 3: Acquisitions processing.

The programs for these tasks and the machine configurations to run them, both in demonstration and in regional center operation, are discussed in the following sections. Task 1 is already in progress and is scheduled for completion by August 1, 1967.

1.1 STUDY PROGRESS AND OBJECTIVES

The magnitude of the services to be provided ultimately by the regional center dictated that the system be partitioned into separate and successive tasks. Three tasks were identified, and each of these was required to meet the simultaneous constraints of effective service to the intended users, economic feasibility, and state-of-the-art program techniques and machine configurations.

To determine the user requirements, three joint meetings were held with the librarians of the six New England state universities, members of their professional staffs, NEBHE representatives and members of our senior staff. Interim reports and working papers [1,2,3] were prepared for these meetings, and were discussed, evaluated, and criticized by all concerned with the aim of deriving a system plan responsive to the needs of all participants. These meetings were of great value in guiding the project.

A systems team visited each of the participating libraries during November and December, and examined their overall operations with emphasis on cataloging and acquisition procedures. The

1. L. F. Buckland, et al. Notes Prepared for Project Meeting on a Regional Library Technical Processing Center, Inforonics, Inc., August 15, 1966.

2. _____, Creation and Initial Use of a Machine Form Catalog of New England Libraries' Holdings by Means of Encoding Current Acquisitions and Collections Being Reclassified, Inforonics, Inc., September 14, 1966.

3. _____, Design of a Technical Processing Center for New England Libraries: Interim Report on Catalog Data Creation, Catalog Data Searching, and Acquisitions Processing, Inforonics, Inc., November 1, 1966.

libraries and staff members were extremely helpful and provided valuable data and insights that will continue to be of use in successive project tasks.

An "overlap" experiment was conducted to determine the degree of collection duplication between libraries. The results are discussed in Appendix 1. Among these, it was shown that a random volume from one library had a 40% chance of being present in another randomly selected library. When current imprint samples were tested, the figure rose to 47%. Rhode Island's holdings were shown to be the "core" collection of greatest duplication elsewhere.

Visits to the Library of Congress were made to determine the status and format of the MARC Project[1], and to discuss the problems of subject authority maintenance and filing rules for machine-form catalogs. A meeting of the USASI subcommittee 15 (Filing) was also attended to discuss the possible recommendations of that group. Our final format decisions were to maintain 100% compatibility with MARC, though in forms more convenient for data manipulation. Appendix 2 discusses this in detail.

The problem of uniform use of L.C. classification schedule is considered in Appendix 3. While the system design of Task 1 makes local exceptions possible, these are not recommended.

During the final stages of this project, we have been studying the machine possibilities for the regional center, and have discussed both existing and soon-to-exist computers, mass storage devices, line printers, and other peripheral equipment with a variety of manufacturers. While computers are experiencing rapid state-of-the-art advances and decreasing costs, progress in mass storage devices has been discouragingly slow by comparison. The configurations we describe are based on equipment existing at present.

---

1. Library of Congress, Information Systems Office A Preliminary Report on the MARC Pilot Project, Washington, D.C., October 1966

2. OVERALL PLAN AND DEVELOPMENT TIMETABLE

The development of the regional center has two component schedules: one for program development and one for performance of service operations. These are shown in Figure 2-1. The program developments have been identified as Tasks 1, 2, and 3, and their descriptions are included in this section. The demonstration of service operations is identified as Task 4 and is also described in this section. Task 1 is currently in progress. Its implementation is described in the Task 1 progress report of March 10, 1967. The system configurations of Tasks 2 and 3 are described in subsequent sections.

Task 1 programming is scheduled for completion on August 1, 1967. The final two months of this task are devoted to checkout and testing, so that Task 2 could begin on July 1, 1967. This is a four month task that includes one month of checkout, so that Task 3 could begin October 1, 1967. Task 3 is a nine month effort that would extend to July 1968.

The successive service operations are shown to begin at the completion of the respective programming tasks. File creation from MARC tapes would begin in August 1967, data file searching in December 1967, and starting in May 1968, a two month transition period would be used to switch over from the magnetic tape based system, adequate for Tasks 1 and 2, to the mass storage based system required for Task 3. In July 1968, all three task programs would be in service operation on the mass storage system.

2.1 TASK 1 SYSTEM: CATALOG DATA FILE CREATION

The function to be performed by Task 1 of this experiment is to develop programs to create a file of catalog data in machine readable form. If the creation of this file is to be economical, it must be a by-product of the individual library's normal operation. One source of the data for the catalog data file (hereafter called CDF) is, therefore, the catalog data created in processing new acquisitions for the library. A second source is Project MARC. A third source is catalog data from any reclassification project contemplated or underway in any of the libraries.

The CDF will be recorded in a format common to all participants and will be compatible with other readily available data. This requirement is best satisfied by using the Library of Congress format as the standard--in particular, its Project MARC format.

2.1.1 Services Provided by Task 1 Programs

Task 1 programs are described in the Task 1 report.[1] The services to the libraries provided by Task 1 programs will be the production of catalog card sets, labels for book spines, and labels

1. The Development of a System for Catalog Data File Creation for the New England Regional Library Technical Processing Center, Inforonics, Inc., March 10, 1967.

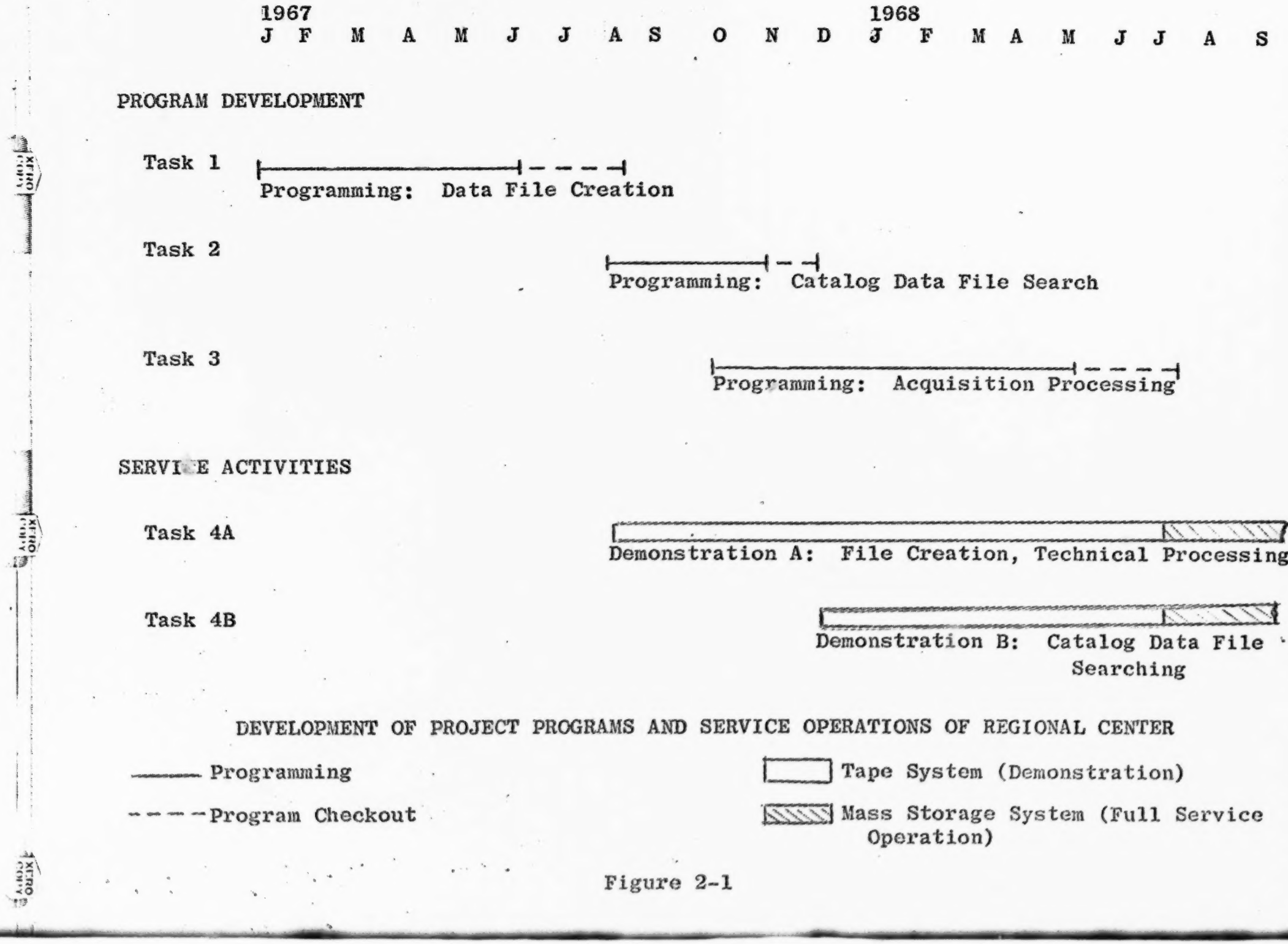
1967
J F M A M J J A S O N D
1968
J F M A M J J A S
PROGRAM DEVELOPMENT
Task 1
Programming: Data File Creation
Task 2
Programming: Catalog Data File Search
Task 3
Programming: Acquisition Processing
SERVICE ACTIVITIES
Task 4A
Demonstration A: File Creation, Technical Processing
Task 4B
Demonstration B: Catalog Data File Searching
DEVELOPMENT OF PROJECT PROGRAMS AND SERVICE OPERATIONS OF REGIONAL CENTER
Programming
Program Checkout
Tape System (Demonstration)
Mass Storage System (Full Service Operation)

Figure 2-1

for book pockets. Catalog card sets will be printed on standard rag stock in upper and lower case. In addition to building the CDF, this will provide immediately useful products for new cataloging and reclassification. When other libraries process the same title, additional card sets can be produced. All card sets will be produced complete with overtyped headings ready for filing.

Book Labels will be produced automatically on a tape typewriter using a Selin platen. Special labeling in addition to the call number will be derived from copy number and book location information which is provided with the catalog copy. The computer will type the copy number and location information in the proper format on the label.

Book pockets will be labeled using pressure sensitive adhesive labels. This service will be optional so that libraries that do not need printed pockets will not receive them.

2.1.2 Use of Task 1 Services in Cataloging New Titles

Catalog card sets for new titles received by the libraries will be produced by the system. When a library catalogs a title, they will request catalog cards and labels for the title by submitting the Library of Congress card number, their location and copy information, and their call number if they do not wish to use the Library of Congress call number. If the record for the item is in the NEBHE file, catalog cards and labels will be produced and shipped to the library. If the record for the item is not in the NEBHE file, they will be informed that the item is not on file. The library will then submit catalog copy, their own original cataloging or Library of Congress catalog copy, to the. center. The center will encode it, type it, enter it into the CDF, and produce and ship the catalog cards to the library. In Task 2, the libraries may request catalog cards by author and/or title as well as by Library of Congress card number.

2.1.3 Use of Task 1 Services for Reclassification

The programs developed in Task 1 could be used to generate catalog cards, shelf list cards, and labels for spines and pockets in a reclassification project. The catalog cards could be generated in alphabetical order to facilitate filing in the catalog. The labels could be generated in Dewey order so that they might easily be matched with the books on the shelves. In addition to providing the proper catalog cards and labels for the items reclassified, a machine readable record for the items reclassified would also be produced. This record could later be used to produce book catalogs or special subject bibliographies or might be used as the basis for record conversion in a mechanized circulation system.

Depending upon the desires of the libraries, the reclassification might proceed class by class, each book in the Dewey class being reclassified into its Library of Congress equivalent, or it

might be done selectively; if the selection were based on date, it could be done by center personnel.

Catalog copy for conversion would be supplied to the center in some form. Micro-filming the entire shelf list would probably be the most efficient means for providing catalog copy in any large scale reclassification effort. When the shelf list does not contain Library of Congress catalog copy, the National Union Catalog and the Library of Congress printed catalogs could be searched if so desired by the libraries. Searching for Library of Congress catalog copy would increase the costs of the reclassification project but would assure that Library of Congress catalog copy would be present in the CDF when it existed for the title. This should appeal to the other libraries who would be using the data in the CDF. The desirability of using a shelf list with a high percentage of Library of Congress cataloging copy as the conversion source is obvious. When Library of Congress cataloging copy cannot be found for the title, the center will input the catalog data as supplied by the library assigning a NEBHE number to the record in lieu of the Library of Congress card number.

The processing center, as it receives reclassification cataloging copy, will encode and key the data in a MARC compatible format, and after proofreading and correcting, will store it in the computer's memory. The center will, on a batch basis, produce catalog cards once a month, shipping the required number of card sets and labels to the originating library. All the libraries will receive notification of all titles reclassified during the month in the form of lists by Library of Congress card number, author, and title.

The other libraries may then request catalog card sets and labels for the items desired by submitting either the Library of Congress card number or the NEBHE identification number plus their own location and copy information.

2.2 TASK 2 SYSTEM: CATALOG DATA FILE SEARCHING

Task 2 consists of the development of techniques and programs to permit searching of the master file by elements other than the L.C. card number and by various combinational elements. We are concerned here with searching in support of cataloging, and the search techniques and programs should be optimal with respect to three somewhat opposing objectives: high probability of matching when unimportant variations occur between question and file, low probability of multiple selection when similar bibliographic elements occur for different titles in the file, and small storage requirements with low running time. We are concerned initially with an off line teletypewriter based system, although the techniques developed will also be applicable to any future display terminal system operating on-line.

2.2.1 Proposed Operation

A library cataloging query will contain one or more of the following bibliographic items:

L.C. card number
Author
Title
Date
Selected Items from Imprint
Selected Items from Collation

At least one of the first three items must be present. The cataloger who composes the query will use his judgment to anticipate the difficulty of the search and include only those items required for unambiguous specification.

The cataloger will prepare Teletype queries to be searched by the center. During the initial test operation, successful searches will result in delivery of catalog cards, and a morning report will be transmitted to each library indicating the status of the previous day's queries. In the case of unsuccessful searches the requesting library will then obtain cataloging information by other means or by original cataloging, and such information will be sent to the center for its file.

2.2.2 Regional Center Search Operation

Task 2 (the searching operation just described) will be performed on the computer by a "batch" processing operation. This means that all search requests received during an interval (perhaps one day) would be collected and the file will be searched that night. A complete set of responses would be printed out for verification. In this verification, the queries which are unsuccessful will then be further investigated by center personnel to insure that the matching programs are accurate. The daily processing report will also notify the requesting cataloger of unsuccessful searches so that he may reformulate his search query if this seems advisable. Man-machine interaction of this sort will be greatly facilitated when the system switches to on-line search and response.

2.2.2.1 Batch Processing

In Task 2 batch processing will be used to save costs during the period when the file is small. This method uses low cost magnetic tape files rather than random access memory. The main catalog will be ordered by L.C. card number. Directories containing abbreviated catalog data will be created for author and title searching.

2.2.3 Search Strategy

The relative efficiency of various search strategies will be explored during Task 2. Search by L.C. card number will be unambiguous, but title and author searches will require more than simple matching in cases where partial or multiple equivalence exists between the search form and the file form. Search strategy in Task 2 (search in support of cataloging) will differ from the searches of Task 3 (search in support of acquisitions) in that it is assumed that the question data will be more definite and specific. The Task 2 question will, in general, be made with the book in hand, whereas the Task 3 question will be composed from data in publishers' catalogs, faculty requests, etc.

The output of the searching of a batch of requests should possibly be arranged in two groups, one of which represents those items with high probability of being correct, and the other containing items about which there is some question. The latter group could be reviewed before card sets are automatically produced and mailed. This will be explored.

2.2.3.1 Search by L.C. Card Number

A first search group would be by L.C. card number with a few characters of the title used for automatic verification. Title segments could be included in the search query and matched, or merely made a part of a response in the case of an on-line system. The output of this search could be in three parts:

A. Complete matches.
B. Match L.C. number, no match on title.
C. No match on L.C. number.

The first category, A, obviously is correct and can be sent on to card production automatically. The second two, B and C, should be returned to the library cataloger. In actual operation, category B will be rare. Category C will be frequent initially and will decrease as the file grows and more MARC data becomes available.

2.2.3.2 Title Search

A second type of search is by title, in which the author, the date of publication, or other items from the imprint or collation statements are used as verification information. The responses can be usefully grouped as follows:

A. Complete match.
B. Incomplete match.
C. No match.

The A's can be processed into cards; the B's returned to the cataloger for specification of a refined search. C's will be identified in the morning report to the requesting library, so that the questioner may reformulate the query, further verify his information, or try to obtain cataloging data from another source.

2.2.3.3 Search by Author

A third type of search question will be by personal or corporate author (corporate author will be indicated in the catalog data as it is in MARC). In addition to author, a few characters of the title and the edition information may be encoded for verification. The response from this type of search will be the following groups:

A. Complete match (one response).
B. Complete match except edition (one response).
C. Complete match with more than one catalog entry.
D. No match.

There is a high probability that A will be correct, and it can be processed further into card sets. Group B will have to be inspected to see if the different edition alters the content of the catalog data. If it is a different edition, the request can be processed further with the different edition statement included in the data, or alternately, L.C. copy for the requested edition can be obtained and keyed. Group C will have to be inspected further to select proper entry. Group D will be entered in the next search procedure.

2.2.3.4 Conference Proceedings

Because of the difficulty of knowing the exact form of the main entry in conference proceedings, these searches may require additional search tools.

The cataloger will search locally using an aid prepared by the center. This searching aid shall be a permuted index of all such proceedings. A permuted index of corporated names may also prove useful.

2.2.4 Use of MARC Data

The MARC data will be searched in a manner identical to that described above. If MARC data enters the system after an entry is already cataloged, it will replace that data already in the file. The libraries will be notified and, if they desire, can order new card sets containing the L.C. catalog data. This condition will be a rare one, however, if the MARC data is as timely as it promises to be. Our master file will contain two sub-

files: data used by a NEBHE participant at least once (largely MARC data), and MARC data not yet used by NEBHE. This partitioning is an internal matter of optimizing file access time, and in the absence of any substantial reclassification project, the entire file will be principally MARC data.

2.2.5 Computer Programming Requirements (Task 2)

The programming for Task 2 includes creating additional files of match information from elements of the master file, generator programs, searching routines for rapidly scanning the directories, printout and response programs for creating hard-copy finding lists and query status output, and directory maintenance programs. Task 2 will include the following programs.

2.2.5.1 Directory Generator

This program creates directory entries from both existing master files and new cataloging data being entered. Directories will be created for author, title, and L.C. card number.

2.2.5.2 Filing Key Generator

This program generates a filing key for the directory from the bibliographic item corresponding to the list. The program removes case information, accent marks, and other information which cannot be handled by a character-by-character alphabetic sorting sequence. It will not at this stage handle anything more complicated than an ordinary alphabetic sort key. More complex filing keys will be developed later as they appear in MARC data.

2.2.5.3 Directory Search and Printout Program

This program compares search entries with the pertinent directory and records matches. The program prints out matches for verification if there is more than one match, or prints out entries for which there is no match. The successful searches which do match are processed by the Task 1 program to create labels, etc. It is likely that two sets of search and response programs will be needed: one for verification of the search without card production, and one used for unattended search and card production.

2.2.5.4 Directory Merge

The new directory information in the master file will be periodically merged to form updated directory files.

2.2.6 System Configuration for Task 2

Task 2 involves creation of a computer system to solve centrally the intellectual workload of cataloging, using a computer service center as a centralizing focus. The mode of processing in this proposed shared cataloging center will be "batch processing". That is to say input requests arrive into the center in a random manner via teletype from the six (or more) participating libraries. Requests are stored into an ordered queue; but the order of the queue is based on the bibliographic content of the request and not the time at which the request arrived (see Task 3). At some arbitrary point in time (at the end of an hour, or two hours, or a day), all the requests accumulated in the queue or batch thus far are processed against the center's master file of member holdings and MARC cataloged data. During the processing a new queue may be beginning to be formed.

Matching the accumulated request queue, now called a request batch, produces actions in accordance with normal find or don't find conditions encountered in a manual search. If the catalog data is available, a set of cards, book labels, etc. is generated for printing and an affirmative answer to the search request is sent over TTY to the requesting library. If the results of the search are null, the requestor is so notified and asked to resubmit the search or to enter the catalog data into the system when available.

The data to be printed is pre-sorted and line printed off-line to the computer (i.e. concurrent with other processing), and then delivered to the appropriate requestor.

In the above, the following main hardware/software functions are implied:

	Function	Software	Hardware
1.	Message Processing	TTY & queue building programs	TTY's; message send/ receive hardware
2.	Batching	Sorting program	4 tape units
3.	Retrieval	Retrieval program	Main computer
4.	Off-line printing	None	Off-line printer

Two classes of computers would be adequate to perform the hardware/ software functions outlined above: medium-scale (MS) and large-scale

(LS). Hardware prices for two low-priced typical systems are given below:

Hardware	Medium Scale	Large Scale
Main computer (8-k memory)	$38,000 (24-k char)	$110,000 (48-k char)
Message Answering	$30,390	$7,500
TTY (6)	$680 (per month)	$680
Tape controller	$10,000	$18,000
4 Tape units	$48,000	$48,000
Off-line printer	$56,375	$56,375
Purchase	$182,765	$239,875
Monthly Rental (= 1/40 Purchase)	$ 4,569	$ 5,408
+ TTY rental	680	680
	$ 5,249	$ 6,088

The differences between these two classes of computers lies in the word size (3 characters per word or 6 characters per word), and the power and sophistication of the instruction set. It should be kept in mind that the larger computer might be easier to adapt the requirements of Task 3.

2.3 TASK 3 SYSTEM: ACQUISITIONS, SEARCH, AND CONTROL

In Task 3 the catalog search capability will be expanded to support acquisitions. In order to support acquisitions search, a large part of the library's complete catalog must be encoded into the catalog data file so that searches can be made against current holdings. This requirement will be satisfied by storing MARC catalog data, encoding current acquisitions and entries from any reclassification process.

The basic requirement of acquisitions searching is the capability for rapid back and forth communication or dialog between the acquisitions searcher and the system. This capability is needed to locate the proper descriptive data about a book given incomplete information. This requirement is satisfied by storing the catalog data in a random access file, instead of magnetic tape, and linking the libraries to a time-shared random access file by a typewriter keyboard.

The primary difference between searching in Task 2 and Task 3 is that in Task 3 the searcher has immediate access to the catalog whereas in Task 2 the searcher has to wait for a batch processing on magnetic tape. This rapid access capability will allow a dialog between the catalog and the searcher, so that the searcher can immediately refine his question if an inquiry fails.

2.3.1 Services to be Provided by Task 3

The following services are to be provided to the library acquisition staffs by Task 3 search capability:

1. Search of holdings.
2. Automatic typing of purchase orders.
3. Maintenance of purchase order control file and automatic typing of claims.
4. Maintenance and automatic typing of want lists.

2.3.1.1 Holdings Search

Requests for file search will be entered in a standard form on the keyboard terminal. The search will consist of L.C. card number or an author or title, a code for requesting library, and vendor name. One of the following responses will be transmitted immediately to the requesting library:

1. Requesting library has item in its collection. Type out notification to searcher.
2. Requesting library has item on order. Type out notification to searcher.
3. Another participating library has item in its collection or on order. This indicates that the central file has the data and that a purchase order will be automatically produced. In addition, a punch card could be issued to be resubmitted to file when requesting library receives item from vendor. The integrated order control system will be designed during the initial period of Task 3.

4. The MARC data file has a record for this item. This is analogous to present manual searching of L.C. proof slip files. Create a purchase order as in 3.

5. No match is found. File searches, which contain incomplete or ambiguous entries, can be processed so that alternative forms of the entry are searched automatically. In addition, the file response will be rapid so if such entry fails to match, the searcher can express it in another form. If all searches fail, then the searcher will have to refer to the conventional manual references.

2.3.1.2 Production of Purchase Order Record

For titles contained in the catalog, the search procedure will create an order record. For searches which yield no response, an order record must be created in machine form. These records will be processed in batches of a suitable size so that all book orders to a vendor may be printed on the same purchase order form. The purchase order form will be produced automatically for mailing to the selected vendor in a window envelope.

2.3.1.3 Production of Control File Information

The order records in machine form will be entered into the center's control file to:

1. Produce a card for the public catalog indicating that the item is now on order. Temporary card will contain basic bibliographic data.

2. Produce a machine record which vendor will enclose in the book. When delivered, it will serve as automatic notification to central file that the item has been received, thereby initiating request for catalog cards and labels.

3. Produce a record for a manual on-order file if this is required.

4. Produce a notification slip to requestor indicating that his request item has been ordered.

5. Encumber departmental or other funds as indicated on purchase request. Fund balance will be updated upon receipt of invoice.

6. Initiate claims when an order remains for a given time period.

2.3.1.4 Book Processing

When books are received, the library will:

1. Alert file when book is received by punch card form returned by vendor with item or by transmitting this information via Teletype terminal.

2. Request catalog card sets with overprinted headings, book labels and pockets using techniques developed in Task 1.

2.3.1.5 Maintenance and Updating of Want Lists to Be Distributed to Antiquarian Dealers

When claim notices are returned notifying the library that a book is not available, this information will be entered into a want list file. This file can be printed out from time to time by language, or country of imprint to facilitate the purchase of books on the out-of-print market. A special notification will be produced for users requesting such books indicating that they are being bought on the out-of-print market.

2.3.2 Detailed Design of an Acquisition System

The acquisitions function performed by libraries are:

(1) Search

- current holdings
- currently on order

(2) Verify entry on request form

- L.C. proof copy file
- other bibliographic tools

(3) Purchase

- select vendor
- produce purchase order
- establish "on-order" file

(4) Receive Item from Vendor

- delete entry in on-order file

- invoice checking

- catalog, process and place book on shelf

Central processing will aid acquisition work by: (1) maintaining a combined holdings file whose entries contain catalog data for an item and the codes for each library holding the item; (2) maintaining L.C. card or MARC data files for verification of the bibliographic data on purchase requests; and (3) allocating storage areas for individual maintenance of on-order files. Economies in this system will be achieved by elimination of manual search and verification procedures, and the carrying over of bibliographic information from acquisition to cataloging.

2.3.3 Preparation of a Request for a Book

The originator of a request for a book will prepare a request form. The request form will by typed and entered into the central system either individually via an on-line typewriter or in batches by a perforated tape. It will contain, in addition to the bibliographic data, the identity of the requesting library and the identity of the vendor.

The items contained in the machine form request will include:

Bibliographic

Main entry (subject to manual verification using standard bibliographic tools, or on-line dialog with control file)

Title

Imprint

Edition

Series

L.C. card number (if available)

Control

Identity of requestor (member of faculty or library staff)

Requestor's address (Dept. - Office no. etc.)

Vendor Code

Identification code for requesting library

The requesting library should use only that amount of information which will insure a response from the file because it will save typing time when entering the request. The L.C. card number, if it is available at the time of the original request as it is in many of the entries in Publisher's Weekly, will be sufficient data for machine file searching.

The purpose of the request will be to determine whether any other member library, including the requesting library, has already ordered the item. This corresponds to a manual search of a card file. The questions to be asked of the file will be:

(1) Does the requesting library have the book on order? (by searching on-order file in allocated area)

(2) Does the requesting library have the book in its collection? (by searching master holdings file for match on selected bibliographic data)

In both cases, the system will produce the appropriate notification for the originator of the request at the individual library.

(3) Do any of the participating libraries have the book in its collection? (after successful search in (2), check holdings code)

If any member holds the item, a notification that catalog cards are available will be issued. The cards will not be issued until notification from the requestor that the book has been received. Delayed release of card sets will eliminate temporary storage of cards and the problem of how to dispose of the cards when an item is not delivered.

2.3.4 Search

2.3.4.1 On-Order Files

In the file storage areas allocated to the individual libraries on-order files will be stored and maintained. This file will be the first step in an acquisitions search. If the search results in a match (an item found), a message to the originator will be issued stating that the item is on order; if no match, the search will proceed to the next step in the acquisitions process.

2.3.4.2 Search Holdings File

A master holdings record will contain a list of all members holding the item requested. If the requested item matches an item in the master holdings file, then one of two conditions exist. Either the library has the item in its collection or some other member has. If it is determined that the item is already in the library's collection, the library will be notified.

If the result is that another member holds the item, the system will produce a set of catalog cards from the stored catalog record upon receipt of the book at the requesting library. The order for a set of catalog cards can be made simultaneously with the decision to order the item. (Matching the bibliographic data also constitutes verification of the entry.)

The on-order file will have its own directories but will point to the catalog data file whenever the requested item is already in this file.

2.3.5 Purchase Order Control

The system will produce a purchase order addressed to the vendor designated in the machine request record. At this time the library's file will be updated to include the ordered item, and a notification that the item is on order will be issued to the originator of the request. The on-order tag in the catalog file will be processed periodically to initiate claims procedures.

2.3.6 Receipt of Item at Originating Library

When the item is received, the entry in the on-order tag will be deleted, catalog card sets will be produced and forwarded to the requestor, and the catalog data file will be updated to include the requestor among the libraries holding the item in their collections.

2.3.7 Use of Common Order-Processing Programs

The functions of book ordering are generally the same from library to library. The variables in the procedure are: the originator of the request, faculty, library staff, or other users; the method of ordering, all books through a single jobber, standing orders to publisher; the format of the request form, format of invoice and bill payment. A central program can be developed which will perform a common set of functions. The input forms on which information items are identified and encoded and the output forms, where printing formats vary, can be controlled by a set of constants

peculiar to each user. In this way, a single program can be used with the result that the programming will be minimum.

2.3.8 Transmittal of Bibliographic Data to Cataloging

When a library originates a request for an item which no other member has in its collection (no entry in the catalog data file) but for which cataloging data is available on the MARC data file, a match resulting from a search of the MARC file indicates verification of the entry. The system will produce a card set when notification of receipt of the item is entered.

If the entry does not exist in any file, a request will be issued so that the requestor can verify the data using other bibliographic tools. When the verification is accomplished at the source, the system will produce the on-order records. When the item is received, a request will be issued to the originator of the order to supply catalog data for the catalog data file.

2.3.9 Computer Programming Requirements (Task 3)

The programming to be done in Task 3 will convert the Task 1 and 2 search procedures to on-line search procedures using a large random access memory.

2.3.9.1 File Creation Program

This program will build the files to be stored in the on-line rapid access memory. These files are the catalog data file, its directories, and the working files used in acquisition. All files will use the same file creation program.

2.3.9.2 File Update Program

This program will accept machine records or data prepared at the typewriter console and will update the files.

2.3.9.3 File Search

This program will find catalog data by searching the finding lists which contain the location of the complete entry.

2.3.9.4 File Printout

This program will produce various specified printouts from the file. For example, for verification an abbreviated printout of an entry can be printed out at the console. When more complete data is required, a full catalog entry can be requested. Printouts of compilations, such as rough bibliographic or reserve book lists, can be produced also.

2.3.9.5 File Transfer

This program is needed to transfer to and from the random access file large blocks of data. These data transfers are needed for safety storage, for updating by batch processing, or for dissemination or distribution of part of the file.

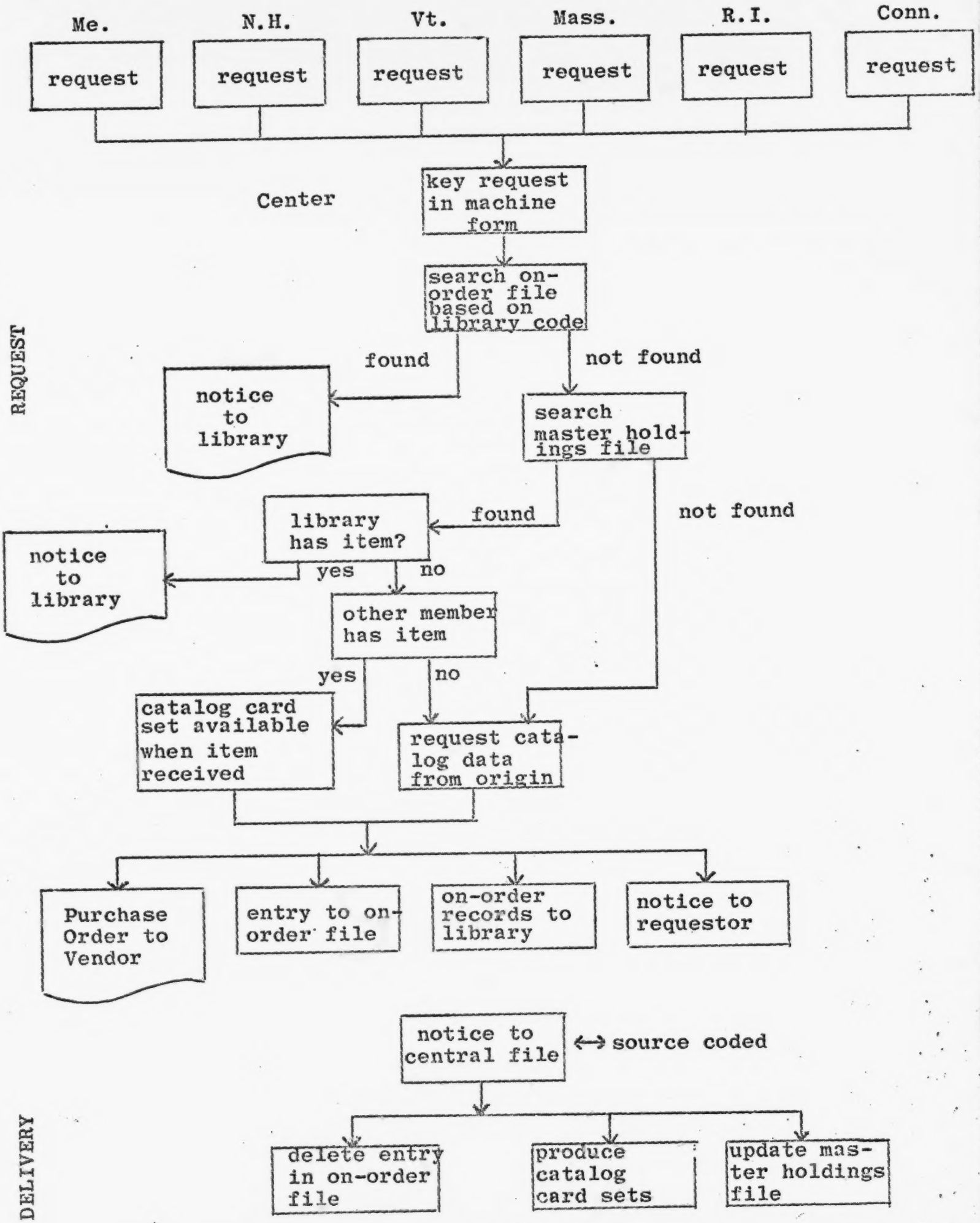

Figure 2-2. Flow Diagram of Acquisitions Support System

XEROX COPY XEROX COPY XEROX COPY

2.3.10 System Configuration and Cost Estimates for Task 3

Task 2 has been described as a shared catalog system, utilizing NEBHE and MARC data, on a batch service basis. Task 3 is a continuation and broadening of this activity in two ways: first, batch processing is replaced by real-time on-line service, and second, processing functions are extended to include more complex bibliographic searches (task 2 search request criteria are limited to LC number, title, author).

In the Task 3 phase of system development, it is anticipated that the data base, in either full or abbreviated form, will be stored on a random access device, on line to the computer. This is in distinction to the serial tape storage techniques proposed for Task 2. The value of random storage (as opposed to serial) is that data can be retrieved from any point in the data storage scheme without having to position a tape forward or backward X hundred feet. Thus access time can be computed in terms of milli-seconds, while tape search time has to be estimated in minutes. This vast increase in speed enables the computer to honor a search request as soon as it is made, perform the search, and return the answer to the requestor within a tolerable waiting period, eliminating the need for batching requests.

This then enables the computer and the requestor to engage in a so-called man-machine dialog, in which the speed and random storage of the computer allows it to respond successive answers to a series of questions asked by a librarian. The question series permits progressive refinement of questions and suggests that more sophisticated bibliographic search requests should be built into the system, exploiting fully the computer's proposed on-line mass data storage.

The hardware costs for Task 3 are: (a) increased core storage to accommodate more sophisticated search programs, and (b) large random access storage device. Prices for these two functions are given below, based on two different scales of computers.

Hardware	Medium Scale	Large Scale
Additional memory (8-k)	$22,000	$49,000
1st Disc memory (146 million char)	$130,400	$158,900
	$152,400	$207,900
Monthly Rental (= 1/40 purchase)	$ 3,810	$ 5,197

As in Task 2, the price difference is accounted for in word size and speed of transfer between the disc and the computer (a large scale 6 character-per-word machine would require half the number of disc-to-memory transfers as a medium scale 3 character-per-word computer). It should be noted that Task 2 service functions will still be ongoing through Task 3. Task 3 might also require additional hardware to accommodate a larger number of users in the network.

2.4 TASK 4: INITIAL USE OF TELETYPE NETWORK AND DEMONSTRATION OF PROGRAMS AND SERVICES

Task 4 will test and demonstrate the programs and services of Tasks 1 and 2 in an environment that will simulate many of the features and problems of the later on-line time-shared system. Remote Teletypes will be installed in each of the six participating libraries and will be connected via Dataphone to a Teletype at the central computer in Cambridge.

During the Task 4 period, all the programs for file creation, technical processing, and catalog data file searching will be used and refined on the basis of this use. Communications procedures will be checked out, an evaluation of the products and services of the system will be made, and time and usage studies will be made to enable a smooth transition to the later on-line system.

The installation of the first Teletypes will begin on August 1, 1967, and after a shakedown period, a three day per week service operation is planned in which the member libraries will request technical processing and search services during the day on Monday, Wednesday, and Friday. Batch processing will be performed the following nights, and technical processing products and answers will be delivered the following day. A principal objective of the Task 4 demonstration system will be to build up the capability for effective and fast response to user requests in preparation for a full service on-line system.

2.5 THE FUTURE NEW ENGLAND LIBRARY NETWORK

It is expected that the transitiion to an on-line time-shared system can be done by July 1968. The first on-line system would provide a full service center for technical processing and acquisitions. A mass storage system and a more powerful time-sharing computer would take the place of the off-line magnetic tape system used in the Task 4 system. The effective operation of the cooperative center will depend on having fast access to a sizeable data base, and the planned transition will provide this

means. The overall plan has been to develop a full repertoire of programs, formats, communication techniques, and methods on an inexpensive magnetic tape system, and then switch over quickly to a full service system. The economic objective is to minimize the initial non-productive set-up time of the on-line system.

Once the on-line system is in operation, plans for expanding the network and its services will be formulated. As the six participating state university libraries develop their own computer facilities, it is reasonable to expect that they will become sub-centers of library networks within their states. Decentralization of this sort will minimize communications costs and optimize the use of facilities. It is probable that improved on-line terminals, incorporating display capability, can soon be added inexpensively at the libraries permitting a greatly expanded search capability. As the Library of Congress MARC services expand, these will be incorporated into the NEBHE system. The early operation of the NEBHE network is expected to provide much data of value in determining the design, operation, and costs of other regional library networks, national library networks, and general machine-aided remote access to information resources.

APPENDIX 1

The Overlap Experiment*

Introduction

As part of a design study to specify a regional processing center for the New England state university libraries, it was required to know the percentage of collection overlap with respect to each of the 30 ordered pairs of 6 libraries. Participating states were Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, and Connecticut. This overlap data was needed to predict the degree of joint use of cataloging information and to estimate the efficiency of collective reclassification.

The results revealed a high degree of commonality in the 6 collections, showing, for example, that a random title from one library had a 40% chance of being present in another randomly selected library. When current imprint samples were tested, the figure rose to 47%. Rhode Island's holdings were shown to be the collection of greatest duplication elsewhere.

Procedure

Two samples of 550 catalog cards were drawn from each library. Sample 1 was an unbiased sample from the general collection; sample 2 was current imprint cards only (1964, 1965, and 1966) The two separate samples were used to determine whether current acquisitions were more or less homogeneous than earlier holdings. The number 550 was chosen to insure that the empirical percentages of overlap were within 5 per cent of their true value with a 99% confidence level. This is explained in detail later. Only monographs subject to L.C. classification were included. Microforms, goverment documents, etc., were excluded as were rare book collections, children's literature, and similar material unlikely to be included in an LC-oriented cooperative system.

Main entry cards were selected to avoid the effects of subject or class bias. The number 550 was divided by the number of drawers in the catalog, and this number, rounded up, was the number taken from each drawer. The distance between successive selections was kept approximately constant, and if a chosen card was inappropriate, successive cards were examined until a selection was made. Sample 1 and sample 2 were drawn separately; the cards were duplicated for distribution to the other libraries and refiled. Selection, reproduction, and refiling time averaged 40 man-hours per sample.

---

*This appendix has been submitted for publication with the title "Statistics of Collection Overlap at the Libraries of the Six New England State Universities".

Each library selected 2 samples, and received 10 samples to be matched. A total of 60 samples was ultimately checked, yielding the overlap for general collections and current imprints for each ordered pair of libraries. Checking time ranged from 6 to 12 man-hours per sample.

Results

Table 1 shows the percentage of overlap from library to library for the two samples. The overlap of sample 1, the general collection, ranged from 28.1% to 55.2%, and had an average of 39.7%. The current imprint sample, sample 2, had a greater range (25.1% to 70.6%), and had a higher average (46.9%). Table 2 lists the percentage probabilities of duplication in each of the libraries, and this appears directly related to library size, as might be expected, but with a shorter range of variation than size alone would cause. This is probably indicative of a "core" collection, common to all, which raises the lower bound of duplication in the smaller libraries. This is further evidenced in table 3 which lists the expected number of duplications elsewhere of titles in a given library. Table 3 contains the row sums of table 1. For the general collection, the average is 1.99 duplications, and 2.35 duplications for current imprints. Approximate library size is shown in table 4.

The increased use of "standing orders" by all participants no doubt accounts in part for this difference. Rhode Island emerges as the most duplicated library, with 2.40 duplicates for the general collection and 2.72 duplicates for the current imprints. This high degree of duplication will result in more efficient use of shared mass storage in the regional center and indicates a high return on cooperative reclassification efforts. The expected number of uses of shared information is the number of duplications plus one, or about 3 uses for an item in the general collection and 3.35 uses if current imprints only are considered.

Appendix: Overlap Probabilities

The overlap percentages of Table 1 may be interpreted as the conditional probability that a title in library i will exist, given that the title exists in library j, the card source, which we will express as:

$$P(l_i | l_j)$$

The probability C that a selected library, l_i, will contain a volume randomly selected from a different library is the sum of the conditional probabilities over j; divided by the number of libraries less 1.

$$C = \sum_{\substack{j \\ l \neq j}} (l_i | l_j) \Big/ \qquad (L-1)$$

These results are shown in Table 2, and are the column sums of the entries in Table 1, divided by 5.

The expected number, D, of duplications, i.e., the expected number of libraries that will have a title contained in a given library, j, will be the sum of the conditional probabilities involving library j.

$$D = \sum_{\substack{i \\ i \neq j}} P(l_i | l_j)$$

For purposes of the experiment, multiple duplications in one library are not considered differently than one duplication. These results are shown in Table 3, and are the row sums of the entries in Table 1. The expected number of uses is the number of duplications plus one.

To obtain the basic conditional probabilities we take the ratio of the k matches found in library i that correspond to titles in the n samples from library j:

$$P(l_i | l_j) = \frac{k_{ij}}{n}$$

The number n is constant for all libraries, and is chosen to be large enough to assure that the discrepancy, δ, between the actual percentage p and the measured percentage p', is less than a chosen amount with a chosen confidence level P. A reasonable requirement for our study was to insure that our estimate overlap percentage was within 5% of the actual overlap percentage, with a confidence level of 99%. That is:

$$P\left\{|p'-p| \leq 0.05\right\} \geq 0.99$$

We use the normal approximation to the binomial distribution and first find the root, x, of the normal distribution for P = .99, yielding x = 2.326. Then using the form:

$$\left(\frac{n}{pq}\right)^{\frac{1}{2}} \geq \frac{x}{\delta} \; ; \quad n \geq pq \frac{x^2}{\delta^2}$$

We assume the worst case (requiring the largest n) for which p = .5, so that:

$$n \geq .25\left(\frac{x}{\delta}\right)^2$$

For the case considered, with x = 2.326 and δ = .05, we find:

$$n \geq 541$$

Were we to reduce the confidence level to 95% and maintain a maximum error of 5%:

$$n \geq 271$$

The former figure is used, rounded up to 550, since we have no previous data to use as verification.

Card Source		Checked by: Maine	N.H.	Vermont	Mass.	R.I.	Conn.
Maine	S1	--	39.7	37.8	39.9	34.9 (3.8)	45.8 (3.5)
	S2	--	42.3	40.6	65.0	41.3 (3.8)	52.4 (6.1)
N.H.	S1	47.0	--	28.1	41.7	34.4 (4.7)	45.6 (3.3)
	S2	43.0	--	36.0	67.6	42.5 (1.7)	39.5 (9.1)
Vermont	S1	43.1	32.9	--	48.8	39.5 (3.4)	49.1 (4.3)
	S2	40.0	55.5	--	70.6	49.6 (0.4)	56.3 (3.9)
Mass.	S1	39.4	29.7	31.6	--	31.1 (2.7)	43.0 (6.0)
	S2	27.9	36.7	25.1	--	28.4 (4.3)	45.0 (3.4)
R.I.	S1	44.7	48.2	39.5	52.2	--	55.2
	S2	43.8	54.9	40.8	69.0	--	63.6 (5.1)
Conn.	S1	35.7	32.5	29.0	41.2	29.8 (4.4)	--
	S2	35.9	48.2	38.6	64.7	43.3 (4.2)	--

PERCENTAGE OVERLAP OF COLLECTIONS

S1 = Sample 1, General Collection
S2 = Sample 2, Current Imprints only (1964-65-66)

Parenthetical figures are additional percentages of imperfect matches (different editions, etc. separately noted by R.I. and Conn., considered as non-match by others).

TABLE 1

Maine	S1	42.0
	S2	38.1
N.H.	S1	36.6
	S2	49.5
Vermont	S1	33.2
	S2	34.3
Mass.	S1	46.8
]	S2	67.4
R.I.	S1	33.9
	S2	41.0
Conn.	S1	47.7
	S2	51.4

Average Percentage Probability of Duplication

General Collection	39.7%
Current Imprint	46.9%

Percentage Probability that a Title in Another Library will be Duplicated in the Listed Libraries

S1 = sample 1, General Collection

S2 = sample 2, Current Imprints Only (1964-65-66)

TABLE 2

Maine	S1	1.98
	S2	2.43
N.H.	S1	1.97
	S2	2.29
Vermont	S1	2.13
	S2	2.72
Mass.	S1	1.75
	S2	1.65
R.I.	S1	2.40
	S2	2.72
Conn.	S1	1.68
	S2	2.30

Average Expected Number of Duplications (excluding multiple copies)

General Collection: 1.99

Current Imprint: 2.35

Average Expected Number of Uses of Shared Cataloging Information (excluding multiple copies)

General Collection: 2.99

Current Imprint: 3.35

Expected Number of Duplications in other Libraries of a Title in the Listed Libraries (excluding multiple copies)

S1 = Sample 1, General Collection

S2 = Sample 2, Current Imprints Only (1964-65-66)

TABLE 3

Maine	450,000
New Hampshire	400,000
Vermont	390,000
Massachusetts	450,000
Rhode Island	275,000
Connecticut	620,000

Number of Volumes Held in 1966 (Approx.)

TABLE 4

APPENDIX 2

NEBHE Compatibility with MARC

1. Introduction

The large main file of the NEBHE Library Facility will have input from three sources: (1) MARC data from the Library of Congress; (2) original cataloging data from current cataloging and (3) cataloging data from any future reclassification project of the NEBHE participants. There can conceivably be other sources, but these three will be considered the only sources initially.

Since MARC data will be one of the primary inputs to the NEBHE Master File, it was decided to insure complete compatibility with MARC. The NEBHE Master File must be derivable from the MARC files. Conversely, if the NEBHE file should be communicated to some other processing center at some future time, the medium of communication would probably be through the MARC format. It is desirable, therefore, to be able to generate a MARC record from the NEBHE Master File. In order to see how this can be done, let us examine the two critical features of the files: (1) file format and (2) character sets.

2. File Formats

The present MARC format contains 25 fixed fields and 20 variable fields. Five of the 25 fixed fields are generated by the computer, one is a blank field, and the remaining 19 are input manually.

The NEBHE format is somewhat similar to MARC. Of the 25 fixed fields, 14 are retained as fixed fields, 3 are retained as modifiers on variable field tags, one is retained as part of variable field data, and 7 are not retained. Five of the 7 that are not retained can be generated automatically whenever desired. One of them, local use, is a blank field. The remaining one, control indicator, is of interest to the Library of Congress only. One of the variable fields in MARC, copy statement, is not retained in NEBHE since it is relevant only to the Library of Congress, The other variable field data are retained without change.

When the Library of Congress converts to the MARC II format, the NEBHE format will be adjusted so that there will again be complete compatibility with the Library of Congress. The presentation that follows lists each one of the MARC fields, describes it, and then gives its treatment in the NEBHE format.

MARC Fixed Field 1, Block Length

This 4 byte fixed field will contain the block length. (It will always be the record length plus four.) The first two bytes specify, in 16-bit binary form, the length of the block, and the next two bytes are blank. The maximum block length is 2008 characters.

NEBHE Format:

This field is not retained in the NEBHE file. It can be generated by computer from the data in file whenever desired.

MARC Fixed Field 2. Record Length

This 4 byte fixed field will contain the record length. The first two bytes specify, in 16-bit binary form, the length of the record, and the next two bytes are blank. The maximum record length is 2004 characters.

NEBHE Format:

This field is not retained in the NEBHE file. It can be generated by computer from the data in file whenever desired.

MARC Fixed Field 3. Library of Congress Catalog Card Number

The LC Card Number will be carried as an eleven-character fixed field which allows for three leading alpha characters and eight numerics. Alpha characters represent any prefix to the LC Card Number. The prefix is left justified with blank fill, and if no prefix is present, three blanks will appear before the numeric card number. The numeric part of the LC Card Number can be described as an eight-digit number; the first two digits are date, the last six digits an identification number. The date will always appear in character positions 4 and 5 of the LC Card Number field. The identification number will be right justified with zero fill. Example:

Number on printed card	Number on magnetic tape
A66-011	Ab1b166000011
AB66-0111	ABb166000111

Note: In the above examples "bl" stands for blank and represents one character position.

NEBHE Format:

This field becomes part of variable field 94 (Library of Congress Catalog Card Number).

MARC Fixed Field 4. Supplement Number

Since supplements, indexes and other dashed-on entries are to be carried as independent records in the MARC Pilot Project, it is necessary to provide a one-character fixed field (numeric) to identify the supplement number. This is required because the supplements will have the same LC Card Number as the original work. If the record does not represent a supplement, this field will be blank.

NEBHE Format:

This field becomes part of fixed field 32.

MARC Fixed Field 5. Type of Main Entry

A one-character fixed field (alpha) will identify the type of main entry as specified in the ALA Cataloging Rules for Author and Title Entries, as follows:

A - Personal Author
B - Government Body
C - Society or Institution
D - Religious Society or Institution
E - Miscellaneous Corporate Body
F - Uniform
G - Title

NEBHE Format:

This field becomes the modifier for the variable fields 10 and 20 (Main Entry and Title Statement). A-F modifies field 10 (Main Entry). G modifies field 20 (Title Statement) if it is the main entry.

MARC Fixed Field 6. Form of Work

A one-character fixed field (alpha) will identify the form of the work. For the MARC Pilot Project, only two forms will be identified.

M - Monograph
S - Serial

NEBHE Format:

This field becomes part of fixed field 32.

MARC Fixed Field 7. Bibliography Indicator

If the work contains any kind of bibliographical apparatus or is itself a bibliography, this fixed field will contain an X. Otherwise, the field will be blank.

NEBHE Format:

This field becomes part of fixed field 32.

MARC Fixed Field 8. Illustration Indicator

If the work contains any type of illustrations other than maps, this fixed field will contain an X. Otherwise, the field will be blank.

NEBHE Format:

This field becomes the modifier for variable field 40 (Collation Statement). If only illustrations are present, the modifier is "i". If both illustrations and maps are present, the modifier is "b".

MARC Fixed Field 9. Map Indicator

If the work contains maps, this fixed field will contain an X. Otherwise, the field will be blank.

NEBHE Format:

This field becomes the modifier for variable field 40 (Collation Statement). If only maps are present, the modifier is "m". If both maps and illustrations are present, the modifier is "b".

MARC Fixed Field 10. Conference or Meeting Indicator

If the work contains the proceedings or the report of a conference, meeting or symposium, this fixed field will contain an X. Otherwise, the field will be blank.

NEBHE Format:

This field becomes part of fixed field 32.

MARC Fixed Field 11. Juvenile Indicator

If the work is for juveniles (as indicated, e.g., by the forms of the subject heading and/or the classification number), this field will contain an X. All records in the Annotated Card program will also contain an X in this field. Otherwise, the field will be blank.

Note that in addition to data in the format of the standard LC printed card, data for material in the Annotated Card program will also be included in the MARC Pilot Project. Thus if a work is in the Annotated Card program, it may be represented in the MARC tape twice: once by the data in standard LC card format, and once by the data in Annotated Card format. Annotated Cards do not have the same LC card number as their counterparts in the regular card program.

Annotated Cards have the prefix AC in the card number and have been designed by the Library of Congress for a card catalog of children's literature. All book publications for children (preschool to young adult) and adult books that appear on young adult and school reading lists will be included in the AC program. Reference books which are suitable for children's use and books about children's literature are also included.

The bibliographic record for the AC card is basically the same as the standard LC card. The only differences are:

a. The card provides a short, clear annotation that will appear as a note.

b. Subject headings specially designed for children will replace the LC subject headings.

c. An Abridged Dewey (9th edition) Classification Number is provided.

d. An "E" for easy book and "FIC" for fiction are provided in place of the Dewey Number when appropriate.

NEBHE Format:

This field becomes part of fixed field 32.

MARC Fixed Field 12. Language Indicator

NEBHE Format:

This field becomes part of fixed field 41.

MARC Fixed Field 13. Language 1

NEBHE Format:

This field becomes part of fixed field 41.

MARC Fixed Field 14. Language 2

The character (alpha) that appears in the Language Indicator describes the use of languages in the work and determines the content of the two language fields as shown below. The codes used in the language fields will be either three or four characters. If three characters, the language code will be left justified with the fourth character position in the field a blank.

S - The work contains only one language. The language will be given in the Language 1 field (alpha) and the Language 2 field will be blank.

T - The work is a translation. Language 1 will contain the language of publication, the Language 2 field (alpha) will contain the language in which the work was originally written or Multilingual, as in the case of anthologies.

M - The work contains more than one language. The principal language will be given in the Language 1 field. If only two languages are used, the second language will be given in the Language 2 field. If more than 2 languages are used, Multilingual will be given as the second language.

D - The work is a dictionary of more than one language. If only 2 languages are used, these will be given in the language fields. If more than 2 languages are used, the principal language will appear in the Language 1 field and Multilingual will be given as the second language.

G - The work is a grammar or reader of the type used in language courses. The native language of the student will be in the Language 1 field. The language being studied will be given in the Language 2 field.

NEBHE Format:

This field becomes part of fixed field 41.

MARC Fixed Field 15. Type of Publication Date

NEBHE Format:

This field becomes part of fixed field 29.

MARC Fixed Field 16. Date 1

NEBHE Format:

This field becomes part of fixed field 29.

MARC Fixed Field 17. Date 2

The contents of the two Date of Publication fields (numeric) are determined by the character that appears in the Type of Publication Date field (alpha) as follows:

S - The date of publication consists of a known date or a probable date that can be represented by four digits, e.g. [1966?]. The date is given in the Date 1 field. The Date 2 field will contain the date of copyright if it appears in the Imprint Statement.

R - The work is a reproduction (such as a reprint or facsimile.) The publication date of the reproduction will be given in the Date 1 field. The Date 2 field will contain the date of original publication.

N - The date of publication is not known. Both date fields will be blank.

M - The date of publication consists of a multiple date. The initial date will be given in the Date 1 field. If the terminal date is known, it will be given in the Date 2 field, otherwise the Date 2 field will be set to the year 9999 to indicate an open ended situation.

Q - One or more of the digits in the imprint date is missing. Such dates will be input as follows:

	Key	Date 1	Date 2
18--	Q	1800	1899
189-	Q	1890	1899

NEBHE Format:

This field becomes part of fixed field 29.

MARC Fixed Field 18. Place of Publication

A four-character fixed field (alpha) will give the place of publication in a four-character mnemonic code.

NEBHE Format:

This field becomes part of fixed field 29.

MARC Fixed Field 19. Publisher

A four-character fixed field (alpha) will contain a publisher code. This code may be two, three, or four characters in length. Should the code contain only two or three characters, it will be left justified within the field, and the remaining character positions in the field will be set to blank(s).

NEBHE Format:

This field becomes part of fixed field 29.

MARC Fixed Field 20. Height of Volume

A two-character fixed field (numeric) represents the height of the volume in centimeters.

NEBHE Format:

This field becomes part of fixed field 29.

MARC Fixed Field 21. Types of Secondary Entries

Each character in this fixed field will indicate a type of secondary entry traced in the record. If the condition is present, the position will contain an X; if not the position will be blank. The fields will be set by the computer programs in the processing of variable field information. The first seven characters described in the following list refer to secondary entries that are not subject added entries.

An X in this position indicates the presence of the name of at least one Personal Author.

An X in this position indicates the presence of the name of at least one Government Body.

An X in this position indicates the presence of the name of at least one Society or Institution.

An X in this position indicates the presence of the name of at least one Religious Society or Institution.

An X in this position indicates the presence of the name of at least one Miscellaneous Corporate Body.

An X in this position indicates the presence of at least one Uniform Heading.

An X in this position indicates the presence of at least one Title secondary entry.

An X in this position indicates the presence of at least one Subject Heading.

NEBHE Format:

This field is not retained in the NEBHE file. It can be generated by computer from the data in file whenever desired.

MARC Fixed Field 22. Series Indicator

If the work is part of a series, this fixed field will contain an X. Otherwise, the field will be blank. This indicator will be set by the computer programs any time a series note is encountered in the variable field processing.

NEBHE Format:

This field is not retained in the NEBHE file. It can be generated by computer from the data in file whenever desired.

MARC Fixed Field 23. Local Use

These character positions have been set aside for the use of the participating libraries.

NEBHE Format:

This field is not retained in the NEBHE file.

MARC Fixed Field 24. Control Indicator

This one-character fixed field (alpha) will be used to indicate the following:

N - This record is new this week.

O - This record was new last week.

R - This record has been revised this week.

(blank) - This record is at least two weeks old.

NEBHE Format:

This field is not retained in the NEBHE file.

MARC Fixed Field 25. Length of Record

This four-character fixed field (numeric) will indicate the total number of characters in the record beginning with the first character of the LC Card Number and ending with the last character of the variable field.

NEBHE Format:

This field is not retained in the NEBHE file. It can be generated by computer from the data in file whenever desired.

MARC Variable Field 1. Main Entry (Tag 10)

A. The format of the main entry will be as specified in the ALA Cataloging Rules for Author and Title Entries subject to the following exceptions for personal names:

(1) Titles (Sir, Lord, etc.) will follow the forename rather than the surname.

Example: Scott, Walter, Sir, bart., 1771-1832.

(2) Date modifiers (b., d., fl., etc.) will follow rather than precede the date(s).

(3) No spaces between initials will be left for future additions

B. For a Title Main Entry, no 10 Tag data will be present in the record. The Title Statement (Tag 20) will be used as the Main Entry.

C. A special character, the pound sign (#), will be used as a delimiter in:

(1) Personal Names

Personal names fall into the following pattern: Name, Title, Date and Relator. The pound sign will be used as follows:

Name
Name,# Date#
Name,# Date#, Relator
Name,# Title
Name,# Title, Date#
Name,# Title, Date# Relator
Name,# Title, Relator
Name,# Relator

After a name, the delimiter follows the punctuation. After a date, the delimiter immediately follows the fourth character of the date.

Example: Churchill, Winston Leonard Spencer,# Sir, 1874#-1965.

Smi[illegible], John, # 1859#, d.

In the above sche[illegible] the word title is used to mean all titles designati[illegible] rank, office or nobility or words or phrases associate[illegible] ith the name. The word relator means those phrases which [illegible]escribe the relationship between an author and a work, e.g., ed., tr., comp., etc.

(2) All other names and headings used as Main Entry

A delimiter will be inserted following the last character that would normally appear in bold face type on an LC printed card.

Example: U. S.# Library of Congress.

California Institute of Technology,# Pasadena.

Note: Some of these delimiters are included in the MARC record for purposes of experimenting at the Library of Congress with computer typesetting, sorting, searching, etc.

NEBHE Format:

The same as in MARC except that this tag is modified as follows:

10a - Personal Author
10b - Government Body
10c - Society or Institution
10d - Religious Society or Institution
10e - Miscellaneous Corporate Body
10f - Uniform

MARC Variable Field 2. Conventional or Filing Title (Tag 15)

The contents of this field will reflect filing procedures currently in use at the Library of Congress.

NEBHE Format:

No change.

MARC Variable Field 3. Title Statement (Tag 20)

This field will contain the Title and all information up to but not including the Edition Statement. A special character delimiter, the pound sign (#), will be used to define the end of the Title and the end of a Short Title, should one exist.

NEBHE Format:

The same as in MARC except that the tag is modified by a "g" if it is a title main entry.

MARC Variable Field 4. Edition Statement (Tag 25)

This field will contain the complete Edition Statement. A special character delimiter, the pound sign (#), will be used to separate edition information from the remainder of the statement.

NEBHE Format:

No change.

MARC Variable Field 5. Imprint Statement (Tag 30)

This field will contain the Imprint Statement and the price of the work if it has been cataloged under the National Program for Acquisition and Cataloging. A special character delimiter, the pound sign (#), will be used to separate place, publisher, date and price. The delimiter will follow the punctuation mark after each subfield when such a mark is present. The presence of only two delimiters indicates the absence of a publisher in the Imprint Statement.

Example: Place,# Publisher,# Date#
Place,# Publisher,# Date# Price
Place,# Date# Price
Place,# Date#

In the above scheme "Place" includes "n. p." and "Date" includes "n. d." Thus: n. p.,# n. d.#

NEBHE Format:

No change.

MARC Variable Field 6. Collation Statement (Tag 40)

This field will contain the Collation Statement.

NEBHE Format:

The same as MARC except that this tag is modified as follows:

40i - Work contains illustrations.
40m - Work contains maps.
40b - Work contains both illustrations and maps.

MARC Variable Field 7. Series Note (Tag 50)

Each series, that is to be traced in exactly the same form as in the Series Note, will be defined by a 50 Tag. Should the Series Note consist of an author and a title, a special character delimiter, the dollar sign ($), will define the end of the author element and the beginning of the title element. This allows the author to appear on one line and the title on another for the overprinting of the series added entry. Tag 50 may be repeated as often as necessary.

The first series note to appear in the record, whether defined by a 50 or 51 Tag, will be the one to be placed in parentheses following the collation statement. Fields tagged 50 or 51 may appear in any sequence, i.e. 50, 51; 51, 50 or any combination for as many Series Notes as appear in the record.

NEBHE Format:

No change.

MARC Variable Field 8. Series Note (Tag 51)

Each series, that is not to be traced in the same form as in the Series Note or that is not to be traced at all, will be defined by a 51 Tag. Tag 51 may be repeated as often as necessary.

The first series note to appear in the record, whether defined by a 50 or 51 Tag, will be the one to be placed in parentheses following the collation statement. Fields tagged 50 or 51 may appear in any sequence, i.e. 50, 51; 51, 50 or any combination for as many Series Notes as appear in the record.

NEBHE Format:

No change.

MARC Variable Field 9. Notes (Tag 60)

This field will contain the information for a Note. Each Note will be found as a separate variable field in the record. Tag 60 may be repeated as often as necessary.

NEBHE Format:

No Change.

MARC Variable Field 10. Subject Tracing (Tag 70)

Each Subject Tracing will be located in a variable field defined by a 70 Tag. Personal names used for Subject Tracings will be formatted in the same manner as for main entries; however, no delimiters will be used. Tag 70 may be repeated as often as necessary.

Should any of the tracings be a 2 line tracing (such as an author/title tracing), a special character delimiter, a dollar sign ($), will define the end of the first element and beginning of the second for the overprinting of the added entry.

NEBHE Format:

No change.

MARC Variable Field 11. Personal Author Tracing (Tag 71)

Each Personal Author Tracing (other than subject) will be located in a variable field defined by a 71 Tag. The format and the use of delimiters for the tracing will follow the same rules that apply to Personal Author Main Entries. Tag 71 may be repeated as often as necessary.

Should any of the tracings be a 2 line tracing (such as an author/ title tracing), a special character delimiter, a dollar sign ($), will define the end of the first element and beginning of the second for the overprinting of the added entry.

NEBHE Format:

No change.

MARC Variable Field 12. Corporate Author Tracing (Tag 72)

Each Corporate Author Tracing (other than subject) will be defined by a 72 Tag. The third character in the tag field will define the 4 types of corporate authors as follows:

B - Government Body
C - Society or Institution
D - Religious Society or Institution
E - Miscellaneous Corporate Body

There are no pound signs (#) used as delimiters. Tag 72 may be repeated as often as necessary.

Should any of the tracings be a 2 line tracing (such as an author/ title tracing), a special character delimiter, a dollar sign ($), will define the end of the first element and beginning of the second for the overprinting of the added entry.

NEBHE Format:

No change.

MARC Variable Field 13. Uniform Tracing (Tag 73)

Each uniform tracing will be defined by a Tag 73. There are no pound signs (#) used as delimiters. Tag 73 may be repeated as often as necessary.

NEBHE Format:

No change.

MARC Variable Field 14. Title Tracing (Tag 74)

Each Title Tracing will be defined by a Tag 74. If the Title Tracing is not the same as the Full Title or Short Title in the Title Statement (Tag 20), it will be given in full in this

field. If a Full Title or a Short Title is to be used from the Title Statement (Tag 20) this field will contain only a "T". There are no pound signs (#) used as delimiters. Tag 74 may be repeated as often as necessary.

NEBHE Format:

The same as MARC except that the "t" is a modifier for Tag 74 rather than part of the data field.

MARC Variable Field 15. Series Tracing (Tag 75)

Each Series Tracing will be defined by a Tag 75. The first character of the data will be either "A" or "T". "A" indicates a Series Tracing consisting of an author and a title. "T" indicates a Series Tracing consisting of a title only. Tag 75 may be repeated as often as necessary.

If the content of the Series Tracing has not been defined by a Tag 50 Series Note, this field will contain either 75A or 75T and the content of the Series Tracing. Should the content of the Series Tracing consist of an author/title, a special character delimiter, the dollar sign ($), will define the end of the author element and the beginning of the title element. This allows the author to appear on one line and the title on another for the overprinting of the series added entry.

If the content of the Series Tracing has been defined by a Tag 50 Series Note, this field will only contain either a 75A or a 75T.

NEBHE Format:

The same as MARC except that the "A" and the "T" are modifiers for Tag 75 rather than part of the data file.

MARC Variable Field 16. Copy Statement (Tag 80)

This field contains information relevant to the Library of Congress.

NEBHE Format:

This field is not retained in the NEBHE file.

MARC Variable Field 17. National Bibliography Number (Tag 83)

A National Bibliography Number is the item number of the publication in the national bibliography. The third character of this tag must be either a one or a zero (blank is invalid). A zero indicates that the national bibliography item number can be accommodated by 15 characters or less. A one indicates that the national bibliography item number requires more than 15 characters. This will usually occur when the national bibliography item numbers for a multi-volume

work are listed. Variable field 83 will contain the item number for the volume number first listed on the Input Work Sheet. The entire series of item numbers for the volumes listed on the Input Work Sheet will be given as the first note tagged 60.

NEBHE Format:

The same as MARC except that "zero" and "1" in the third character position of the Tag are replaced by "a" and "b" respectively.

MARC Variable Field 18. Library of Congress Call Number (Tag 90)

This field will contain a complete Library of Congress Call Number or a Library of Congress Class Number or it will be blank.

The presence of only a Class Number indicates that the Library of Congress cataloged the publication but has not added it to its collections and does not expect to do so. On the standard Library of Congress printed card such Class Numbers are enclosed within brackets. In the MARC Pilot Project such Class Numbers will not be enclosed within brackets but instead the Class Number will be followed by the letters NLC, which stand for "Not in LC". Two spaces will separate the letters NLC from the Class Number.

The absence of a Library of Congress Call Number indicates that the publication has been assigned to the Law Library of the Library of Congress.

NEBHE Format:

The same as MARC except that Tag 90 is modified by an "n" if call number contains class numbers only.

MARC Variable Field 19. Dewey Decimal Classification Number (Tag 92)

This field will contain the complete Dewey Decimal Classification Number. Some Dewey Numbers are prefixed by a "j". When such a number appears in the MARC Pilot Project, the "j" will follow the Dewey Number separated by two spaces and will be enclosed within parentheses. The "j" is used for juvenile titles and the numbers have been assigned from the 9th abridged edition of the Decimal Classification. Those cards in the Annotated Card program may have an "E" for easy book and "FIC" for fiction provided in place of the Dewey Number when appropriate.

NEBHE Format:

No change.

MARC Variable Field 20. Library of Congress Catalog Card Number (Tag

If the LC Card Number has date and number (66-1037) or a prefix with a date and number (AC66-1037), the entire card number (prefix and numerics) will appear in the fixed field. Whenever the LC Card Number includes a suffix (66-1037/CD), the suffix will not appear in the fixed field for the LC Card Number but rather will appear with Tag 94 in the variable field. The suffix field has a maximum character length of three. If there is no suffix to the card number, no Tag 94 will appear in the variable field. Since less than 2 percent of the catalog cards contain suffixes and the suffix does not affect the uniqueness of the number, the decision was made not to include the suffix in the fixed field.

NEBHE Format:

The entire LC Card Number--prefix, numerics, and and suffix--is placed in one field, variable field Tag 94. In the case of original cataloging, the NEBHE processing center will assign a NEBHE card number for this field. In this case Tag 94 will be modified as follows, depending on where the original cataloging is done:

c - Connecticut
e - Maine
m - Massachusetts
n - New Hampshire
r - Rhode Island
v - Vermont

3. Character Sets

While the NEBHE keyboard is not yet firm, there are two distinct probabilities at this time. First, during the test phase of the NEBHE project, copy will be prepared at the computer center using a Dura Mach 10. During later phases of the project when the various libraries go on-line with teletypes, the seven level ASCII code will probably be used. A table comparing these character sets is given here:

MARC	NEBHE (Dura)	NEBHE (ASCII)
1. Upper and lower case alphabetics	1. Upper and lower case alphabetics	1. Upper and lower case alphabetics
2. Numerics	2. Numerics	2. Numerics
3. Punctuation and special char.:	3. Punctuation and special char.:	3. Punctuation and special char.:
comma	comma	comma

MARC	NEBHE (Dura)	NEBHE (ASCII)
period	period	period
semicolon	semicolon	semicolon
hyphen	hyphen	hyphen
left bracket	left bracket	left bracket
right bracket	right bracket	right bracket
left paren	left paren	left paren
right paren	right paren	right paren
colon	colon	colon
slash	slash	slash
ampersand	ampersand	ampersand
question mark	question mark	question mark
dollar	hollow square	dollar
sharp	solid triangle	sharp
single quote	times	single quote
double quote	en dash	double quote
underscore	em dash	asterick
	equals	equals
	plus	plus
	exclamation	exclamation
	percent	percent
	less than	less than
	greater than	greater than
	quote	at sign
	unquote	backwards slant
		up arrow
		left arrow
4. Diacritics	4. Diacritics	4. Diacritics
breve	angstrom	(none)
umlaut		
acute		
cedilla		
angstrom		
macron		
tilde		
grave		
circumflex		
5. Special:	5. Special:	5. Special:
space	space	space
tab	tab	tab
carriage return	carriage return	carriage return
upper case	upper case	upper case
lower case	lower case	lower case
delete	delete	delete
backspace	backspace	line feed
stop code	stop code	vertical tab
non-print		form feed
print restore		reader on

skip restore
end of fold
end of records
end of tape

reader off
tape on
tape off
acknowledge
alternate mode
horizontal tab
end of transmis-
sion

It is planned to make up the diacritics in the NEBHE Master File through the use of an escape code □ . Thus □ ap will represent an umlaut, for example. Thus the NEBHE system will have a translator to go from MARC to NEBHE and vice versa.

APPENDIX 3

LIBRARY OF CONGRESS CLASSIFICATION SCHEDULES AND THE NEBHE PROJECT

The Library of Congress Classification Schedules are made by the Library of Congress especially for use by its staff to classify books for its own collections. Libraries using these schedules to classify books for their own collections often disagree with where and how certain subjects are handled by L.C. and attempt to modify the schedules to varying degrees. This can have serious economic and processing results.

Since L.C. cards are available for a large percentage of the books cataloged for any library (with the likelihood under Title II that an even larger percentage will be available in the future) and all carry L.C. class numbers as assigned by L.C. catalogers, there is a great saving in being able to use these numbers without change or review. Deviations from the schedules as printed will require that every book acquired by a library be reviewed and possibly changed to meet local preference. This is expensive as it requires the review to be done by professional librarians.

Therefore, careful consideration should be given to any proposed deviation before a decision is made to include it or not. The library considering making deviations must satisfactorily answer the question of whether the proposed deviation will be worth the increased cost and effort. In general, it has been found that local deviations cannot be economically justified, and in reality, accomplish little. A list of the most commonly used deviations is included below together with recommendations as to how they should be handled in the NEBHE Project.

L.C. Class No.	Comment
AC	Various kinds of sets and series are classed here, subarranged by language. Use this number for general collections which do not show a subject bias. Those that do show a subject bias, class with the subject.
AP	Periodicals (General). Subarrangement is by language, then by place. An alternative widely accepted is straight arrangement by title. Deviation by title is minor and should be considered if deemed important.

L.C. Class No.	Comment
B	There is much dissatisfaction with this schedule, and there are many deviations.
BF	Psychology. Often put with psychiatry in R. Libraries should follow L.C. and use BF.
BL-BX	Religion. Large specialized collections (such as Union Theological) do not use this classification at all. Depending upon the religion of the library, there are alternate schedules (not L.C.) which are substituted. Deviation for religion recommended only for libraries having large special collections. The typical university library should use these numbers without change.
CT	Biography. Many libraries class all biography in CT rather than to distribute subject biography with subject. L.C. practice should be followed strictly here, that is, use CT only for general biography, class subject oriented biography with subject.
GN	Anthropology. Some libraries have preferred to class certain materials belonging here with the corresponding history numbers in D through F. General university libraries should follow L.C. practice and class all anthropology here.
HA	Statistics. To be used only for general statistics. Some libraries have classed all statistical material here. Special statistics should be placed with subject as is done by L.C.
HX	Socialism, Communism, Anarchism, Bolshevism. Obviously a weak L.C. number. Some libraries prefer to class certain materials in political science (J) or history (D-F). In general, HX should be used, though certain exceptions may be made.

L.C. Class No.	Comment
J	Political Science. Schedule is badly out of date. L.C. is at this point being strongly urged to revise schedule. Libraries should follow L.C. practice until revision is completed. UN publications which are classed here are an especial problem. For economy and simplicity, follow L.C. again.
JX	International Law. Number has been badly abused because of the absence of a K schedule. K schedule is presently being published. Each item should be individually considered, and older materials classed to reflect current practice.
K	Currently being published. Where actual schedules have been constructed, use them. Where no schedule exists do one of the following: For general law: Use K and author cutter number knowing that this group of numbers will have to be reclassed in the future. Use K and a country subdivision followed by an author cutter, reclass when schedules are available. Class law of a particular subject or industry with that subject (e.g., space law in Q or T).
LT	Textbooks. Often classed with subject rather than with LT. Use this number only for general textbooks.
LH-LJ	University, college and school magazine, and college fraternities and their publications. Often placed in LD through LF. Probably should be classed as per L.C.
M	Music. Certain deviations made. Unless schedule is being used for a large, special collection, M should be used as it is written.

L.C. Class No.	Comment
N	Fine arts. Same comments as for M.
P	Literature and language. Main difficulty is PZ3-4. L.C. is studying the problem and will probably offer alternatives. Libraries should class fiction with original distinguishing English language translations via the floating tables. PZ5 is often used for ALL juvenile literature. Unless a library has a large collection of same, they should follow L.C. practice rather than to class all juvenile materials here.
QD	Chemistry. Libraries often adapt Dewey convention for the arrangement of the elements. Libraries should follow L.C. convention of arranging alphabetically.
QP	Physiology. Libraries have preferred R and BF for certain materials belonging here. Unless library is strictly medical or psychology, follow L.C. practice.
QR	Bacteriology. Often classed in R. Unless library is a special medical library, use QR.
R	Special medical libraries often substituted special classifications for R (i.e., NLM, etc.) Unless collection is for a medical library, use R as printed.
RC	Psychiatry. Distinction between this number and BF arbitrary. Some libraries place much BF here. For general libraries, prefer L.C. practice.
SK	Hunting sports. Often combined with GV. Not recommended. Follow L.C. practice closely.
T	Technology. Distinction between Q and T often arbitrary. Follow L.C. practice closely. Problem of technology and economics (HD9000-9999) often cause libraries to class materials in the HD number. Follow L.C. practice closely.

L.C. Class No.	Comment
	Bibliography. One schedule most deviated from. L.C. is studying problem and possibly will revise schedule or provide alternate placement. Most libraries prefer to place bibliography with subject. This is a major administrative decision. If Z is abandoned, and bibliography is placed with subject, some scheme should be devised so that bibliographies can be retrieved at some future date and reclassified according to L.C.'s anticipated revision. The following alternates have been used: Class bibliography of a subject with subject in the general/special number if there is one, or with the general number for the subject. Class with general number of a subject and cutter .A1 (followed by author number) or .Z6 (followed by author number). Class with general number for a subject and add mnenomic .016 to class number. Any of the schemes work. The last (using .016) has the one advantage that it can be more easily machine retrieved.